R E M A R K S

A non-final Office Action was mailed on October 17, 2007. Claims 1-9 are pending.

Background

By way of background, and before addressing the prior art rejections, Applicant wishes to reiterate, as set forth in the specification between pages 3-4, that the speech control unit comprises a keyword recognition system for recognition of a predetermined keyword that is spoken by the user and which is represented by a particular audio signal and the speech control unit being arranged to control the beam forming module, on basis of the recognition of the predetermined keyword, in order to enhance second components of the audio signals which represent a subsequent utterance originating from a second orientation of the user relative to the microphone array. ***The keyword recognition system is arranged to discriminate between audio signals related to utterances representing the predetermined keyword and to other utterances which do not represent the predetermined keyword.*** The speech control unit is arranged to re-calibrate if it receives sound corresponding to the predetermined keyword, from a different orientation. Preferably this sound has been generated by the user who initiated an attention span (see also Fig. 3) of the apparatus to be controlled. There will be no re-calibration if the predetermined keyword has not been recognized. As a consequence, ***speech items spoken from another orientation and which are not preceded by the predetermined keyword, will be discarded***.

Thus, the predetermined keyword is used to calibrate the beam forming module to a particular user's utterances and orientations and to recognize speech of a user who is moving in an environment in which other users might speak too. (Applicant's specification, page 4, lines 29-31) The predetermined keyword is used to discriminate between users subject to the recognition system, and other users.

This background discussion is only presented to summarize different aspects of Applicant's disclosure, and should not be considered as limiting the claims in any way.

Response to Arguments

Claims 1-9 are rejected under 35 U.S.C. §103(a) as being unpatentable over Schroder et al. (U.S. Patent 7,136,817) in view of Kaufholz. It is believe that "Kaufholz" refers to U.S. Patent 7,050,971, although such reference is not listed on the Notice of References Cited PTO-892 accompanying the Office Action of October 17, 2007. The Examiner is respectfully requested to confirm this by citing the appropriate document on a subsequently-issued PTO-892.

Claim 1 as amended requires, inter alia, *a speech recognition unit for creating an instruction for the apparatus based on recognized speech items of the speech signal, further comprising a **keyword recognition system** for recognition of **a predetermined keyword that is spoken by the user** and which is represented by a particular audio signal and the speech control unit being arranged to control the beam forming module, **on basis of the recognition of the predetermined keyword**, in order to **enhance second components** of the audio signals which represent a subsequent utterance originating from a second orientation of the user relative to the microphone array; wherein the recognition of the predetermined keyword at the first orientation and the subsequent utterance at the second orientation **calibrates** the beam forming module **to follow** the user from the first orientation to the second orientation.*

Thus, the predetermined keyword is used to associate the system with a particular user uttering such keyword, and to follow such user to a different orientation based on subsequent utterances of such user. The predetermined keyword thus **_calibrates_** the system to a particular user such that the system can distinguish between users at different orientations uttering the predetermined keyword, and others.

Schroder et al. teaches a control system based on recognizing and associating speech inputs of a user at one time, with speech inputs from another time. This is further clarified in claim 1 of Schroder et al.:

> *Method for the voice control of a device appertaining to consumer electronics, comprising: converting speech inputs of a user into digitized voice signals; extracting first features, which are characteristic of individual sounds of the speech, and thus permit speech recognition of the spoken sounds, from the*

- 6 -

digitized voice signals; ___converting said extracted first features into voice___

___commands for controlling said device___*; extracting second features, which permit*

a characterization of the voice of the respective user and are used for

distinguishing between the speech inputs of different users, from the digitized

voice signals; and accepting after a voice command from a first user further voice

commands only from this first user, ___by testing whether said second features of___

___further speech inputs corresponding to said further voice commands-can be___

___assigned to the same user___*. (emphasis added)*

In Schroder et al., a comparison is made between "first features" and "second features" of voice

signals as a means for confirming the identity of the user contributing such features. Thus,

discrimination between users is based ___purely on speech inputs___ at different times ___without regard___

___to the content___ of such speech inputs.

Accordingly, one skilled in the art would not consider the claimed invention to be

obvious in view of Schroder et al., alone or in combination with Kaufholz, because Schroder et

al. fails to teach or reasonably suggest the functionality of the claimed **predetermined keyword**

separate from the speech signals associated therewith. In other words, in the system of Schroder

et al., as long as subsequent voice commands defined by subsequent speech inputs are provided

by the same user, through an analysis and comparison of speech inputs only, the subsequent

voice commands will be followed. In the present case, subsequent voice commands will not be

followed through the control of the beam forming module without a threshold determination

made by the existence of the predetermined keyword, as set forth for example in the following

claim element from claim 1, as amended, *"control the beam forming module,* ___on basis of the___

___recognition of the predetermined keyword___*, in order to* ___enhance second components of the___

___audio signals which represent a subsequent utterance___ *originating from a second orientation of*

the user relative to the microphone array."

Furthermore, the predetermined keyword is used to associate the beam forming module

in a manner that is separate and distinct from the teaching in Schroeder et al. and Kaufholz.

Kaufholz simply teaches that a beam former can track an audio source such as by identifying the loudest source:

> The apparatus 220 has two microphone inputs 224 and 226 for receiving the microphone signals from the respective outputs 204 and 214. All microphone signals (in the example two external microphone signals and one internal microphone signal) are supplied to a beam former 240. The beam former combines the microphone signals, resulting in a higher performance and resolution of the resulting microphone signal. **_The beam former may also select or even 'track' an audio source. Typically, the loudest source signal is identified (usually a person speaking) and this source signal is tracked among the various microphone input signals._** The output signal of the beam former is provided to the microphone input 110 of the audio cancellation unit 100. Also shown are two audio inputs 228 and 230 of the apparatus 220 which serve to receive audio signal from respective external apparatuses. In the shown system, the external audio inputs 228 and 230 are connected to the respective audio line outputs 206 and 216 of the audio set 200 and the TV 210. Within the apparatus 220, the external audio inputs 228 and 230 are connected to the respective audio inputs 120 and 130 of the audio cancellation module 100. (Kaufholz, column 5, lines 8-27).

Schroeder et al. also discloses the detection and tracking of voice signals through voice characteristics:

> For the detection of the voice signals, a single microphone or else a microphone array comprising two or more microphones may be provided. The microphone array may, for example, be integrated in a television receiver. The microphones convert the detected sound signals into electric signals, which are amplified by amplifiers, converted by AD converters into digital signals and then fed to a signal processing unit. **_The latter can take into account the respective place where the user is located by a different scaling or processing of the detected_**

> ***sound signals.*** *Furthermore, a correction of the microphone signals with respect to the sound signals emitted from the loudspeakers may also take place. The signal conditioned in this way is then fed to the speech recognition unit and speaker recognition unit, it being possible for algorithms or hardware units to be configured separately or else jointly. The commands determined and the identity of the user are then finally fed to a system manager for controlling the system. (Schroeder et al., column 4, lines 44-62)*
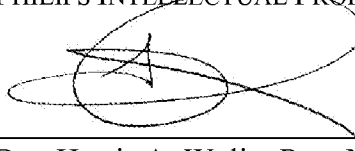
However, neither Schroder et al. nor Kaufholz teach or reasonably suggest the control of a beam forming module through the recognition of a predetermined keyword **as claimed**. Thus, one skilled in the art would not consider the claimed invention to be obvious in view of the combination of Schroeder et al. and Kaufholz.

In view of the above amendments and remarks, it is believed that claims 1-9 are in condition for allowance. However, if for any reason the Examiner should consider this application not to be in condition for allowance, the Examiner is respectfully requested to telephone the undersigned attorney at the number listed below prior to issuing a further Action.

Any fee due with this paper, not already paid through an EFS-Web filing, may be charged to Deposit Account No. 50-3894. Any overpayment may be credited to Deposit Account No. 50-3894.

Respectfully submitted,

PHILIPS INTELLECTUAL PROPERTY & STANDARDS

By: Harris A. Wolin, Reg. No. 39,432
Attorney for the Applicant

Please Address All Correspondence to:
Yan Glickberg, Registration No. 51,742
Phone: (440) 483-3455
CUSTOMER NUMBER 24737

- 9 -